

Practitioner's Docket No. 915-006.037

**PATENT**

**IN THE UNITED STATES PATENT AND TRADEMARK OFFICE**

In re application of: K. HALONEN et al.  
Application No.: 0 To Be Assigned Group No.: To Be Assigned  
Filed: Herewith Examiner: To Be Assigned  
For: Method and Device for Providing Speech-enabled Input  
in an Electronic Device Having a User Interface

Assistant Commissioner for Patents  
Washington, D.C. 20231

**TRANSMITTAL OF CERTIFIED COPY**

Attached please find the certified copy of the foreign application from which priority is claimed for this case:

Country: International/WIPO  
Application Number: PCT/IB03/01262  
Filing Date: April 07, 2003

**WARNING:** "When a document that is required by statute to be certified must be filed, a copy, including a photocopy or facsimile transmission of the certification is not acceptable." 37 C.F.R. § 1.4(h) (emphasis added).

  
SIGNATURE OF PRACTITIONER

Reg. No. 27,550

Alfred A. Fressola

(type or print name of practitioner)

Tel. No. (203) 261-1234

Ware, Fressola, Van Der Sluys & Adolphson LLP

P.O. Address Bradford Green, Building Five  
755 Main Street, P.O. Box 224  
Monroe, CT 06468

Customer No.: 004955

**NOTE:** The claim to priority need be in no special form and may be made by the attorney or agent, if the foreign application is referred to in the oath or declaration, as required by § 1.63.

**CERTIFICATE OF MAILING (37 C.F.R. § 1.8a)**

I hereby certify that this correspondence is, on the date shown below is being deposited with the United States Postal Service with sufficient postage as ~~first class~~ Express mail in an envelope addressed to the Assistant Commissioner for Patents, Washington, D.C. 20231. Express

Express Mail No. EV 435647992UUS

Date: April 07, 2004

  
Signature

Annemarie Maher

(type or print name of person certifying)

(Transmittal of Certified Copy [5-4])



**WORLD INTELLECTUAL PROPERTY ORGANIZATION  
ORGANISATION MONDIALE DE LA PROPRIÉTÉ INTELLECTUELLE**

34, chemin des Colombettes, Case postale 18, CH-1211 Genève 20 (Suisse)  
Téléphone: (41 22) 338 91 11 - e-mail: wipo.mail @ wipo.int. - Fac-similé: (41 22) 733 54 28

**PATENT COOPERATION TREATY (PCT)  
TRAITÉ DE COOPÉRATION EN MATIÈRE DE BREVETS (PCT)**

**CERTIFIED COPY OF THE INTERNATIONAL APPLICATION AS FILED  
AND OF ANY CORRECTIONS THERETO**

**COPIE CERTIFIÉE CONFORME DE LA DEMANDE INTERNATIONALE, TELLE QU'ELLE  
A ÉTÉ DÉPOSÉE, AINSI QUE DE TOUTES CORRECTIONS Y RELATIVES**

International Application No. } PCT/IB 03 / 01262  
Demande internationale n° }

International Filing Date }  
Date du dépôt international }

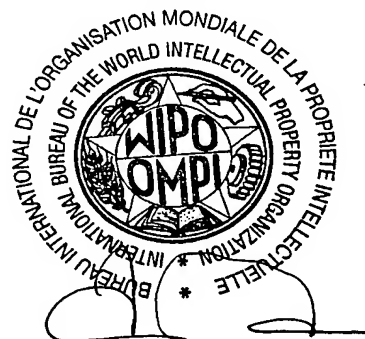
**07 APRIL 2003**  
**(07.03.03)**

Geneva/Genève,

**15 MARCH 2004**  
**(15.03.04)**

**International Bureau of the  
World Intellectual Property Organization (WIPO)**

**Bureau International de l'Organisation Mondiale  
de la Propriété Intellectuelle (OMPI)**



J.-L. Baron

Head, PCT Receiving Office Section  
Chef de la section "office récepteur du PCT"

PCT

## REQUEST

The undersigned requests that the present international application be processed according to the Patent Cooperation Treaty.

For receiving Office use only	
PCT / IB 0 3 / 0 1 2 6 2	
International Application No.	
07 APRIL 2003	(07. 04. 03)
International Filing Date	
INTERNATIONAL BUREAU OF WIPO	
PCT International Application	
Name of receiving Office and "PCT" International Application"	
Applicant's or agent's file reference (if desired) (12 characters maximum) 51507 WO	

Box No. I TITLE OF INVENTION		Method and device for providing speech-enabled input in an electronic device having a user interface	
Box No. II APPLICANT		<input type="checkbox"/> This person is also inventor	
Name and address: (Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country. The country of the address indicated in this Box is the applicant's State (that is, country) of residence if no State of residence is indicated below.)		Telephone No.	
Nokia Corporation Keilalahdentie 4 02150 Espoo FINLAND		Facsimile No.	
		Teleprinter No.	
		Applicant's registration No. with the Office	
State (that is, country) of nationality: Finland		State (that is, country) of residence: Finland	
This person is applicant for the purposes of:		<input type="checkbox"/> all designated States <input checked="" type="checkbox"/> all designated States except the United States of America <input type="checkbox"/> the United States of America only <input type="checkbox"/> the States indicated in the Supplemental Box	
Box No. III FURTHER APPLICANT(S) AND/OR (FURTHER) INVENTOR(S)			
Name and address: (Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country. The country of the address indicated in this Box is the applicant's State (that is, country) of residence if no State of residence is indicated below.)		This person is:	
Halonen Katriina Katajarinne 9 37470 Vesilahti Finland		<input type="checkbox"/> applicant only <input checked="" type="checkbox"/> applicant and inventor <input type="checkbox"/> inventor only (If this check-box is marked, do not fill in below.)	
		Applicant's registration No. with the Office	
State (that is, country) of nationality: Finland		State (that is, country) of residence: Finland	
This person is applicant for the purposes of:		<input type="checkbox"/> all designated States <input type="checkbox"/> all designated States except the United States of America <input checked="" type="checkbox"/> the United States of America only <input type="checkbox"/> the States indicated in the Supplemental Box	
<input checked="" type="checkbox"/> Further applicants and/or (further) inventors are indicated on a continuation sheet.			
Box No. IV AGENT OR COMMON REPRESENTATIVE; OR ADDRESS FOR CORRESPONDENCE			
The person identified below is hereby/has been appointed to act on behalf of the applicant(s) before the competent International Authorities as:		<input checked="" type="checkbox"/> agent <input type="checkbox"/> common representative	
Name and address: (Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country.)		Telephone No.	
Straus, Alexander (Dr.) Becker, Kurig, Straus Bavariastrasse 7 80336 München Germany		0049-89-746 303 0	
		Facsimile No.	
		0049-89-746 303 11	
		Teleprinter No.	
		Agent's registration No. with the Office	
<input type="checkbox"/> Address for correspondence: Mark this check-box where no agent or common representative is/has been appointed and the space above is used instead to indicate a special address to which correspondence should be sent.			

Form PCT/RO/101 (first sheet) (March 2001; reprint January 2003)

See Notes to the request form

BEST AVAILABLE COPY

BEST AVAILABLE COPY

07/04 '03 LUN 16:09 [N° TX/RX 8833]

Sheet No. ...2...

Continuation of Box No. III <b>FURTHER APPLICANT(S) AND/OR (FURTHER) INVENTOR(S)</b>	
If none of the following sub-boxes is used, this sheet should not be included in the request.	
Name and address: (Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country. The country of the address indicated in this Box is the applicant's State (that is, country) of residence if no State of residence is indicated below.)  <b>Salminen, Henri</b> <b>Kankurinmäentie 86 as. 9</b> <b>07280 Iloa</b> <b>Finland</b>	This person is: <input type="checkbox"/> applicant only <input checked="" type="checkbox"/> applicant and inventor <input type="checkbox"/> inventor only (If this check-box is marked, do not fill in below.)  Applicant's registration No. with the Office
State (that is, country) of nationality: <b>Finland</b>	State (that is, country) of residence: <b>Finland</b>
This person is applicant for the purposes of: <input type="checkbox"/> all designated States <input type="checkbox"/> all designated States except the United States of America <input checked="" type="checkbox"/> the United States of America only <input type="checkbox"/> the States indicated in the Supplemental Box	
Name and address: (Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country. The country of the address indicated in this Box is the applicant's State (that is, country) of residence if no State of residence is indicated below.)	This person is: <input type="checkbox"/> applicant only <input type="checkbox"/> applicant and inventor <input type="checkbox"/> inventor only (If this check-box is marked, do not fill in below.)  Applicant's registration No. with the Office
State (that is, country) of nationality:	State (that is, country) of residence:
This person is applicant for the purposes of: <input type="checkbox"/> all designated States <input type="checkbox"/> all designated States except the United States of America <input type="checkbox"/> the United States of America only <input type="checkbox"/> the States indicated in the Supplemental Box	
Name and address: (Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country. The country of the address indicated in this Box is the applicant's State (that is, country) of residence if no State of residence is indicated below.)	This person is: <input type="checkbox"/> applicant only <input type="checkbox"/> applicant and inventor <input type="checkbox"/> inventor only (If this check-box is marked, do not fill in below.)  Applicant's registration No. with the Office
State (that is, country) of nationality:	State (that is, country) of residence:
This person is applicant for the purposes of: <input type="checkbox"/> all designated States <input type="checkbox"/> all designated States except the United States of America <input type="checkbox"/> the United States of America only <input type="checkbox"/> the States indicated in the Supplemental Box	
Name and address: (Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country. The country of the address indicated in this Box is the applicant's State (that is, country) of residence if no State of residence is indicated below.)	This person is: <input type="checkbox"/> applicant only <input type="checkbox"/> applicant and inventor <input type="checkbox"/> inventor only (If this check-box is marked, do not fill in below.)  Applicant's registration No. with the Office
State (that is, country) of nationality:	State (that is, country) of residence:
This person is applicant for the purposes of: <input type="checkbox"/> all designated States <input type="checkbox"/> all designated States except the United States of America <input type="checkbox"/> the United States of America only <input type="checkbox"/> the States indicated in the Supplemental Box	
Name and address: (Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country. The country of the address indicated in this Box is the applicant's State (that is, country) of residence if no State of residence is indicated below.)	This person is: <input type="checkbox"/> applicant only <input type="checkbox"/> applicant and inventor <input type="checkbox"/> inventor only (If this check-box is marked, do not fill in below.)  Applicant's registration No. with the Office
State (that is, country) of nationality:	State (that is, country) of residence:
This person is applicant for the purposes of: <input type="checkbox"/> all designated States <input type="checkbox"/> all designated States except the United States of America <input type="checkbox"/> the United States of America only <input type="checkbox"/> the States indicated in the Supplemental Box	
<input type="checkbox"/> Further applicants and/or (further) inventors are indicated on another continuation sheet.	

Form PCT/RO/101 (continuation sheet) (March 2001; reprint January 2003)

See Notes to the request form

BEST AVAILABLE COPY

Sheet No. ...3...

## Box No. V DESIGNATION OF STATES

Mark the applicable check-boxes below; at least one must be marked.

The following designations are hereby made under Rule 4.9(a):

## Regional Patent

- ☒ AP ARIPO Patent: GH Ghana, GM Gambia, KE Kenya, LS Lesotho, MW Malawi, MZ Mozambique, SD Sudan, SL Sierra Leone, SZ Swaziland, TZ United Republic of Tanzania, UG Uganda, ZM Zambia, ZW Zimbabwe, and any other State which is a Contracting State of the Harare Protocol and of the PCT (if other kind of protection or treatment desired, specify on dotted line) .....
- ☒ EA Eurasian Patent: AM Armenia, AZ Azerbaijan, BY Belarus, KG Kyrgyzstan, KZ Kazakhstan, MD Republic of Moldova, RU Russian Federation, TJ Tajikistan, TM Turkmenistan, and any other State which is a Contracting State of the Eurasian Patent Convention and of the PCT
- ☒ EP European Patent: AT Austria, BE Belgium, BG Bulgaria, CH & LI Switzerland and Liechtenstein, CY Cyprus, CZ Czech Republic, DE Germany, DK Denmark, EE Estonia, ES Spain, FI Finland, FR France, GB United Kingdom, GR Greece, IE Ireland, IT Italy, LU Luxembourg, MC Monaco, NL Netherlands, PT Portugal, SE Sweden, SI Slovenia, SK Slovakia, TR Turkey, and any other State which is a Contracting State of the European Patent Convention and of the PCT
- ☒ OA OAPI Patent: BF Burkina Faso, BJ Benin, CF Central African Republic, CG Congo, CI Côte d'Ivoire, CM Cameroon, GA Gabon, GN Guinea, GQ Equatorial Guinea, GW Guinea-Bissau, ML Mali, MR Mauritania, NE Niger, SN Senegal, TD Chad, TG Togo, and any other State which is a member State of OAPI and a Contracting State of the PCT (if other kind of protection or treatment desired, specify on dotted line) .....

## National Patent (if other kind of protection or treatment desired, specify on dotted line):

- |   |  |   |
|---|--|---|
| <input checked="" type="checkbox"/> AE United Arab Emirates .....               | <input checked="" type="checkbox"/> GM Gambia .....                                    | <input checked="" type="checkbox"/> NZ New Zealand .....                      |
| <input checked="" type="checkbox"/> AG Antigua and Barbuda .....                | <input checked="" type="checkbox"/> HR Croatia .....                                   | <input checked="" type="checkbox"/> OM Oman .....                             |
| <input checked="" type="checkbox"/> AL Albania .....                            | <input checked="" type="checkbox"/> HU Hungary .....                                   | <input checked="" type="checkbox"/> PH Philippines .....                      |
| <input checked="" type="checkbox"/> AM Armenia .....                            | <input checked="" type="checkbox"/> ID Indonesia .....                                 | <input checked="" type="checkbox"/> PL Poland .....                           |
| <input checked="" type="checkbox"/> AT Austria .....                            | <input checked="" type="checkbox"/> IL Israel .....                                    | <input checked="" type="checkbox"/> PT Portugal .....                         |
| <input checked="" type="checkbox"/> AU Australia .....                          | <input checked="" type="checkbox"/> IN India .....                                     | <input checked="" type="checkbox"/> RO Romania .....                          |
| <input checked="" type="checkbox"/> AZ Azerbaijan .....                         | <input checked="" type="checkbox"/> IS Iceland .....                                   | <input checked="" type="checkbox"/> RU Russian Federation .....               |
| <input checked="" type="checkbox"/> BA Bosnia and Herzegovina .....             | <input checked="" type="checkbox"/> JP Japan .....                                     | <input checked="" type="checkbox"/> SC Seychelles .....                       |
| <input checked="" type="checkbox"/> BB Barbados .....                           | <input checked="" type="checkbox"/> KE Kenya .....                                     | <input checked="" type="checkbox"/> SD Sudan .....                            |
| <input checked="" type="checkbox"/> BG Bulgaria .....                           | <input checked="" type="checkbox"/> KG Kyrgyzstan .....                                | <input checked="" type="checkbox"/> SE Sweden .....                           |
| <input checked="" type="checkbox"/> BR Brazil .....                             | <input checked="" type="checkbox"/> KP Democratic People's Republic of Korea .....     | <input checked="" type="checkbox"/> SG Singapore .....                        |
| <input checked="" type="checkbox"/> BY Belarus .....                            | <input checked="" type="checkbox"/> KR Republic of Korea .....                         | <input checked="" type="checkbox"/> SK Slovakia .....                         |
| <input checked="" type="checkbox"/> BZ Belize .....                             | <input checked="" type="checkbox"/> KZ Kazakhstan .....                                | <input checked="" type="checkbox"/> SL Sierra Leone .....                     |
| <input checked="" type="checkbox"/> CA Canada .....                             | <input checked="" type="checkbox"/> LC Saint Lucia .....                               | <input checked="" type="checkbox"/> TJ Tajikistan .....                       |
| <input checked="" type="checkbox"/> CH & LI Switzerland and Liechtenstein ..... | <input checked="" type="checkbox"/> LK Sri Lanka .....                                 | <input checked="" type="checkbox"/> TM Turkmenistan .....                     |
| <input checked="" type="checkbox"/> CN China .....                              | <input checked="" type="checkbox"/> LR Liberia .....                                   | <input checked="" type="checkbox"/> TN Tunisia .....                          |
| <input checked="" type="checkbox"/> CO Colombia .....                           | <input checked="" type="checkbox"/> LS Lesotho .....                                   | <input checked="" type="checkbox"/> TR Turkey .....                           |
| <input checked="" type="checkbox"/> CR Costa Rica .....                         | <input checked="" type="checkbox"/> LT Lithuania .....                                 | <input checked="" type="checkbox"/> TT Trinidad and Tobago .....              |
| <input checked="" type="checkbox"/> CU Cuba .....                               | <input checked="" type="checkbox"/> LU Luxembourg .....                                | <input checked="" type="checkbox"/> TZ United Republic of Tanzania .....      |
| <input checked="" type="checkbox"/> CZ Czech Republic .....                     | <input checked="" type="checkbox"/> LV Latvia .....                                    | <input checked="" type="checkbox"/> UA Ukraine .....                          |
| <input checked="" type="checkbox"/> DE Germany .....                            | <input checked="" type="checkbox"/> MA Morocco .....                                   | <input checked="" type="checkbox"/> UG Uganda .....                           |
| <input checked="" type="checkbox"/> DK Denmark .....                            | <input checked="" type="checkbox"/> MD Republic of Moldova .....                       | <input checked="" type="checkbox"/> US United States of America .....         |
| <input checked="" type="checkbox"/> DM Dominica .....                           | <input checked="" type="checkbox"/> MG Madagascar .....                                | <input checked="" type="checkbox"/> UZ Uzbekistan .....                       |
| <input checked="" type="checkbox"/> DZ Algeria .....                            | <input checked="" type="checkbox"/> MK The former Yugoslav Republic of Macedonia ..... | <input checked="" type="checkbox"/> VC Saint Vincent and the Grenadines ..... |
| <input checked="" type="checkbox"/> EC Ecuador .....                            | <input checked="" type="checkbox"/> MN Mongolia .....                                  | <input checked="" type="checkbox"/> VN Viet Nam .....                         |
| <input checked="" type="checkbox"/> EE Estonia .....                            | <input checked="" type="checkbox"/> MW Malawi .....                                    | <input checked="" type="checkbox"/> YU Yugoslavia .....                       |
| <input checked="" type="checkbox"/> ES Spain .....                              | <input checked="" type="checkbox"/> MX Mexico .....                                    | <input checked="" type="checkbox"/> ZA South Africa .....                     |
| <input checked="" type="checkbox"/> FI Finland .....                            | <input checked="" type="checkbox"/> MZ Mozambique .....                                | <input checked="" type="checkbox"/> ZM Zambia .....                           |
| <input checked="" type="checkbox"/> GB United Kingdom .....                     | <input checked="" type="checkbox"/> NO Norway .....                                    | <input checked="" type="checkbox"/> ZW Zimbabwe .....                         |
| <input checked="" type="checkbox"/> GD Grenada .....                            |  |   |
| <input checked="" type="checkbox"/> GE Georgia .....                            |  |   |
| <input checked="" type="checkbox"/> GH Ghana .....                              |  |   |

Check-boxes below reserved for designating States which have become party to the PCT after issuance of this sheet

- ☒ NI Nicaragua .....
- ☐ .....

**Precautionary Designation Statement:** In addition to the designations made above, the applicant also makes under Rule 4.9(b) all other designations which would be permitted under the PCT except any designation(s) indicated in the Supplemental Box as being excluded from the scope of this statement. The applicant declares that those additional designations are subject to confirmation and that any designation which is not confirmed before the expiration of 15 months from the priority date is to be regarded as withdrawn by the applicant at the expiration of that time limit. (Confirmation (including fees) must reach the receiving Office within the 15-month time limit.)

Sheet No. ...4...

Box No. VI PRIORITY CLAIM				
The priority of the following earlier application(s) is hereby claimed:				
Filing date of earlier application (day/month/year)	Number of earlier application	Where earlier application is:		
		national application: country or Member of WTO	regional application:* regional Office	international application: receiving Office
item (1)				
item (2)				
item (3)				
item (4)				
item (5)				

☐ Further priority claims are indicated in the Supplemental Box.

The receiving Office is requested to prepare and transmit to the International Bureau a certified copy of the earlier application(s) (only if the earlier application was filed with the Office which for the purposes of this international application is the receiving Office) identified above as:

☐ all items ☐ item (1) ☐ item (2) ☐ item (3) ☐ item (4) ☐ item (5) ☐ other, see Supplemental Box

\* Where the earlier application is an ARIPO application, indicate at least one country party to the Paris Convention for the Protection of Industrial Property or one Member of the World Trade Organization for which that earlier application was filed (Rule 4.10(b)(ii)): . . .

---

**Box No. VII INTERNATIONAL SEARCHING AUTHORITY**

Choice of International Searching Authority (ISA) (if two or more International Searching Authorities are competent to carry out the international search, indicate the Authority chosen; the two-letter code may be used):

ISA / SE . . . . .

Request to use results of earlier search; reference to that search (if an earlier search has been carried out by or requested from the International Searching Authority):

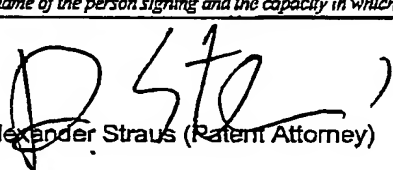
Date (day/month/year) . . . . . Number . . . . . Country (or regional Office) . . . . .

---

**Box No. VIII DECLARATIONS**

The following declarations are contained in Boxes Nos. VIII (i) to (v) (mark the applicable check-boxes below and indicate in the right column the number of each type of declaration):		Number of declarations
<input type="checkbox"/> Box No. VIII (i)	Declaration as to the identity of the inventor	:
<input type="checkbox"/> Box No. VIII (ii)	Declaration as to the applicant's entitlement, as at the international filing date, to apply for and be granted a patent	:
<input type="checkbox"/> Box No. VIII (iii)	Declaration as to the applicant's entitlement, as at the international filing date, to claim the priority of the earlier application	:
<input type="checkbox"/> Box No. VIII (iv)	Declaration of inventorship (only for the purposes of the designation of the United States of America)	:
<input type="checkbox"/> Box No. VIII (v)	Declaration as to non-prejudicial disclosures or exceptions to lack of novelty	:

Sheet No. 5

Box No. IX CHECK LIST; LANGUAGE OF FILING		
<p>This international application contains:</p> <p>(a) in paper form, the following number of sheets:</p> <p>request (including declaration sheets) : 5</p> <p>description (excluding sequence listings and/or tables related thereto) : 17</p> <p>claims : 3</p> <p>abstract : 1</p> <p>drawings : 2</p> <p>Sub-total number of sheets : 28</p> <p>sequence listings : </p> <p>tables related thereto : </p> <p>(for both, actual number of sheets if filed in paper form, whether or not also filed in computer readable form; see (c) below)</p> <p>Total number of sheets : 28</p> <p>(b) <input type="checkbox"/> only in computer readable form (Section 801(a)(i))</p> <p>(i) <input type="checkbox"/> sequence listings</p> <p>(ii) <input type="checkbox"/> tables related thereto</p> <p>(c) <input type="checkbox"/> also in computer readable form (Section 801(a)(ii))</p> <p>(i) <input type="checkbox"/> sequence listings</p> <p>(ii) <input type="checkbox"/> tables related thereto</p> <p>Type and number of carriers (diskette, CD-ROM, CD-R or other) on which are contained the</p> <p><input type="checkbox"/> sequence listings: .....</p> <p><input type="checkbox"/> tables related thereto: .....</p> <p>(additional copies to be indicated under items 9(ii) and/or 10(ii), in right column)</p>	<p>This international application is accompanied by the following item(s) (mark the applicable check-boxes below and indicate in right column the number of each item):</p> <p>1. <input checked="" type="checkbox"/> fee calculation sheet : </p> <p>2. <input type="checkbox"/> original separate power of attorney : </p> <p>3. <input type="checkbox"/> original general power of attorney : </p> <p>4. <input type="checkbox"/> copy of general power of attorney; reference number, if any: ..... : </p> <p>5. <input type="checkbox"/> statement explaining lack of signature : </p> <p>6. <input type="checkbox"/> priority document(s) identified in Box No. VI as item(s): ..... : </p> <p>7. <input type="checkbox"/> translation of international application into (language): ..... : </p> <p>8. <input type="checkbox"/> separate indications concerning deposited microorganism or other biological material : </p> <p>9. <input type="checkbox"/> sequence listings in computer readable form (indicate type and number of carriers)</p> <p>(i) <input type="checkbox"/> copy submitted for the purposes of international search under Rule 13ter only (and not as part of the international application) : </p> <p>(ii) <input type="checkbox"/> (only where check-box (b)(i) or (c)(i) is marked in left column) additional copies including, where applicable, the copy for the purposes of international search under Rule 13ter : </p> <p>(iii) <input type="checkbox"/> together with relevant statement as to the identity of the copy or copies with the sequence listings mentioned in left column : </p> <p>10. <input type="checkbox"/> tables in computer readable form related to sequence listings (indicate type and number of carriers)</p> <p>(i) <input type="checkbox"/> copy submitted for the purposes of international search under Section 802(b-quater) only (and not as part of the international application) : </p> <p>(ii) <input type="checkbox"/> (only where check-box (b)(ii) or (c)(ii) is marked in left column) additional copies including, where applicable, the copy for the purposes of international search under Section 802(b-quater) : </p> <p>(iii) <input type="checkbox"/> together with relevant statement as to the identity of the copy or copies with the tables mentioned in left column : </p> <p>11. <input type="checkbox"/> other (specify): ..... : </p>	Number of items
Figure of the drawings which should accompany the abstract: 1	Language of filing of the international application: English	
<p>Box No. X SIGNATURE OF APPLICANT, AGENT OR COMMON REPRESENTATIVE</p> <p>Next to each signature, indicate the name of the person signing and the capacity in which the person signs (if such capacity is not obvious from reading the request).</p> <p></p> <p>Dr. Alexander Straus (Patent Attorney)</p>		

For receiving Office use only		2. Drawings: <input type="checkbox"/> received:  <input type="checkbox"/> not received:
1. Date of actual receipt of the purported international application: 07 APRIL 2003 (07. 04. 03)	3. Corrected date of actual receipt due to later but timely received papers or drawings completing the purported international application:	
4. Date of timely receipt of the required corrections under PCT Article 11(2):	5. International Searching Authority (if two or more are competent): ISA/SE	6. <input checked="" type="checkbox"/> Transmittal of search copy delayed until search fee is paid
Date of receipt of the record copy by the International Bureau:		

Form PCT/RO/101 (last sheet) (January 2003)

See Notes to the request form

BEST AVAILABLE COPY

07/04 '03 LUN 16:09 [N° TX/RX 8833]

**Method and Device for providing speech-enabled input  
in an electronic device having a user interface**

5

The present invention relates to multimodal interactive browsing on electronic devices and portable terminals and in communication networks. More specifically, the invention relates to a simple multi-modal user interface concept, by offering a close guidance of possible voice data-input and voice browsing as an entry alternative to use manual input. Moreover, the invention is related to checking the preliminary conditions that should be fulfilled for valid voice input.

10

In multimodal applications, users can interact with other input modalities than only the keypad. For example, commands that are traditionally given by scrolling and clicking can be speech-enabled in the application so that the user can speak the commands, which will then be recognized by an automatic speech recognition engine. Adding speech interaction to visual applications receives growing interest as the enabling technologies are maturing, since in many mobile scenarios using the keypad is difficult, for example when driving or walking.

15

Until now different multimodal browsing architectures have already been proposed. For example the document US 6101473 describes a method, where voice browsing is realized by synchronous operation of a telephone network service and an internet service. This is definitively prohibitive due to the waste of network resources, requiring two different communication links. Further this service requires an interconnection between the telephone service and the internet service. Another hurdle for user satisfaction is that the over-the-air co-browser synchronization required in a distributed browser architecture may cause latencies in browser operation which will degrade the user experience.

20

25

The document US 6188985 describes a method in which a wireless control unit implements the voice browsing capabilities to a host computer. For this purpose, a number of multimodal browser architectures have been proposed where these operations are placed on a network server.

30

The patent US 6374226 describes a system that is capable of changing the speech recognition grammar dynamically. For example, when an E-mail program goes to the composition mode, a new grammar set up is dynamically activated. This includes on one hand an improved use of device resources, but also includes the severe disadvantage that the device changes its „passive

35

**BEST AVAILABLE COPY**



vocabulary". This may lead to frustrating experiences as the user who has learned that the device understands a certain expression may be faced with a device feigning deafness for its input when running another application.

- 5 The known systems suffer from the fact that users are not very keen to take the speech-enabled features into use. Another problem arising from the state of the art is that users may not always be aware of the operation status of speech enabled browsing systems.

10 While there are standards being developed for how to write multimodal applications, there are no standards as to how the application interface should be built so that it would be as easy as possible for the user to become aware that speech input can be used.

Especially in devices and applications it would be desirable for a user to know which particular voice input is allowed at different times or under certain conditions.

15

But when a user has put a speech recognition system successfully into use, it is probable that the user also continues to use it. In other words, there is a hurdle in starting to use speech control.

20

The problem has been solved earlier by audio prompts etc., but these become annoying very quickly, which degrades the usability experience.

Moreover, due to system load or the behavior of applications, all speech control options may not be available at all times, which is very difficult to convey to the user using prior art techniques.

25

All the above approaches for a multimodal browsing architecture have in common that they are not suitable for use in mobile electronic devices or terminals such as mobile phones, or handheld computers, due to low computing power, restricted resources or low battery capacity.

30

So it would be desirable to have a multimodal browsing system that is speech-enabled and provides superior user-friendliness.

35

According to a first aspect of the present invention, there is provided a method for multimodal interactive browsing, comprising the steps of activating a multimodal user interaction, comprising at least one key input option and at least one voice input option, displaying the at least one key input option, checking if there is at least one condition affecting said voice input option, and providing voice input options and displaying indications of said provided voice input options according to said condition.

The activation of a multimodal user interaction in which at least one key input option and conditionally at least one voice input option is provided, can be provided by at least switching on the device or by activating a respective menu or respective settings.

5

In said multimodal browsing key input options are unconditionally provided and said at least one voice input option is conditionally provided. Said at least one voice input option is not provided, if at least one condition that could possibly interfere with voice input is fulfilled. The condition can be, e.g. ambient noise or a too low signal to noise ratio at the audio input. The condition can be, e.g. too low processing power, or battery status. The condition can e.g. be a too low voice transfer capability in case of a distributed speech input / recognition system. The condition can be restricted device resources. It should be noted that the conditions affecting the voice recognition feature can be caused by a combination of the above conditions.

10

15 The at least one key input option is displayed on a display of said electronic device or mobile terminal device, as in the case of conventional devices and conventional browsing.

The method is characterized by checking if at least one condition affecting the voice input is fulfilled and providing said at least one voice input option and displaying indications of said voice input options on said display, in case that none of said conditions is fulfilled. The checking can be performed e.g. every second, in faster intervals or continually. The checking can also be performed in an event controlled manner, wherein the check is only performed if an event is detected that is indicative of an impossible voice input.

20

25 If no such condition is fulfilled, the method can provide at least one voice input option and displaying indications of said at least one available voice input option on said display. It is also possible to display, if no such condition is fulfilled the depiction or representation or indication that a voice input option is present and that a voice input can actually be performed. The first part describes the principle that a voice input can be made or is in the passive vocabulary of a voice recognition engine and the second part describes that voice recognition engine is active.

30

It is also possible to display a representation of the checked condition that is actually fulfilled and interferes with the voice input option. This can be embodied as e.g. a kind of icon or text indicating what kind of condition prevents the voice input and how it may be removed.

35

In multimodal applications where voice input can be given in addition to visual input (using keypad), the user must be made aware of when voice inputs are possible and also what is the

allowed input. This method suggests a transparent way of letting the user know exactly when voice recognition is active and which voice commands are speech-enabled at any point.

- 5 Event mechanisms can also be used by the system to determine situations when speech recognition is not available for unexpected reasons or when the application designer has specified that a certain command or a command set is speech-enabled. All commands that are speech-enabled at a certain moment will be marked with a suitable visual method, for example coloring, to indicate to the user both the speaking moment and the allowed utterance.
- 10 The invention proposes to indicate dynamically by visual keywords or visual cues the elements that can be voice controlled depending on the availability of voice control for each item. For example, if the speech recognition engine cannot be used temporarily, or if only certain options are available at a certain point in an application, only those options are highlighted on the screen.
- 15 It can also be marked when speech input is temporally unavailable. It is also possible to mark only entries that are not speech enabled. This is some kind of an inverse approach that can be extended to some kind of switching between marking speech enabled straight and marking not speech enabled input options in dependence of the number of markings necessary. This can be implemented straight forward: green: enabled and black: not enabled and in an inverse notation
- 20 red: not speech enabled input options and black: speech enabled input options.

This invention suggests visual keywords or cues to indicate to the user what can be spoken and also when the speech-enabling is on or off. When a visual command is speech-enabled, the command itself is marked e.g. with a different color or a respective icon than the commands that

25 are not speech-enabled. When the speech-enabling is off, the color or a respective icon of the command changes dynamically back, and if speech-enabling is turned on again, the color or icon will change again. This marking will immediately indicate to the user what can be said and when. The method can be combined with an input prediction method to sort frequently used input option to the top of the list.

30

The reasons why the speech-enabling of a command might change while the user stays on the same screen can be, for example, the following:

- System error: connection to the speech recognizer is cut off unexpectedly,
  - Change of environment: the device detects too much background noise for the recognition
- 35 to work properly,

- System is currently doing some action during which it cannot listen at the same time because of system or application limitations, exhausted or exhaustively used system resources e.g. fetching data for the user, and
- Application designer's choice, described more closely in the following paragraph

5

Different applications may choose different recognition grammars and vocabulary to enable speech in different manners, and the usage can vary even within one application. For example, if on one screen the user can do several different actions (each including 2-3 choices of a menu), the order of which does not matter, it is reasonable to allow the user to speak any of the options. On the next screen, there may again be several actions, but this time the order is not totally free. It is best to guide the user's speech input by making the order of actions explicit with the visual speech-enabling cue that is chosen, highlighting the actions at their proper time.

10

15

However, in a totally eyes-free situation, where voice is the only available modality, this invention cannot be used as the only cue to the user. Some auditory keywords would be required to indicate to the user when (and/or what) the user can speak. One way to indicate that a speech recognition is actually available can be implemented by a vibration alarm prompt. The vibration alarm prompt can comprise a single vibration as a start signal, and a short double vibration as a stop signal.

20

In an example embodiment said displayed indications of voice input options comprise keywords. The keywords can visualize available voice input or control options. The keywords can comprise any kind of cues or hints to the actual speech input that may be not displayable (such as whistling, humming or such sounds).

25

In another example embodiment said displaying of indications of said voice input options on said display further comprises displaying, if a speech recognition is actually possible. As already described above that is the recording or recognition state of a speech or voice recognition engine. This can be described as 'recording' or 'recognizing' sign.

30

In another example embodiment of the present invention said displaying of indications of voice input options comprises displaying said voice input options itself. That is, the input options are depicted as the verbatim of the words to be spoken for the voice input. The wording "input option" has been carefully chosen not to restrict the indication or the input option to any kind of specific form.

35

In another example embodiment of the present invention said displaying of indications of said voice input options on said display, is provided with a hysteresis. The use of a hysteretical behavior helps to avoid fast changes on the indication of the availability of said voice input options, in case that one of said checked conditions is near a threshold between inferring and not inferring said voice input feature. The hysteresis can be implemented in the checking or the program performing the check, or in the application performing the indication.

In another example embodiment of the present invention said displaying of indications of said voice input options on said display, is provided with a backlog function. As in the case of the hysteresis the backlog function can be used to determine and eliminate fast changing conditions that may cross a threshold value related to a condition (e.g. even overriding the hysteresis) to prevent the user from being confused by a rapidly changing voice input ability or voice input options. A backlog functionality can be implemented by a storage for storing the checking results of the last 'n' seconds and a deactivation of a voice input option, as long as a single "over threshold value" entry is present in said back log file. As in the case of the hysteresis, the backlog function can be implemented in the display application or in the checking application. In both cases, the information conveyed to the user is made independent from small changes in the vicinity of a threshold and from fast changes.

According to yet another aspect of the invention, a software tool is provided comprising program code means for carrying out the method of the preceding description when said program product is run on a computer, a network device or a mobile terminal device.

According to another aspect of the present invention, a computer program product downloadable from a server for carrying out the method of the preceding description is provided, which comprises program code means for performing all of the steps of the preceding methods when said program is run on a computer, a network device or a mobile terminal device.

According to yet another aspect of the invention, a computer program product is provided comprising program code means stored on a computer readable medium for carrying out the methods of the preceding description, when said program product is run on a computer, a network device or a mobile terminal device.

According to another aspect of the present invention a computer data signal is provided. The computer data signal is embodied in a carrier wave and represents a program that makes the computer perform the steps of the method contained in the preceding description, when said

computer program is run on a computer, a network device or a mobile terminal device.

The computer program and the computer program product may be distributed in different parts and devices of the network. The computer program and the computer product device run in  
5 different devices e.g. terminal device and remote speech recognition engine of the network. Therefore, the computer program and the computer program device have to be different in abilities and source code.

According to yet another aspect of the present invention a mobile terminal device for executing  
10 simulated communication is provided. The terminal device comprises a central processing unit, a display, a key based input system, a microphone and a data access means.

The central processing unit CPU, is provided to execute and run applications on said mobile terminal. The display is connected to said CPU, to display visual content received from said  
15 CPU. The key based input system is connected to said CPU, to provide a key input feature that can provide key input options displayed on said display. The microphone is connected to said CPU, to provide a conditional voice input feature. The data access means is connected to said CPU, to handle data and to exchange data required for the operation of the CPU. In the simplest case the data access means is a storage and in more sophisticated embodiments the data access  
20 means can comprise e.g. a modem for a network access.

The CPU is configured to perform multimodal browsing via said display, said key based input system and said microphone. The CPU is configured to continually monitor conditions that interfere with said voice input and to provide said voice input feature, and display an indication  
25 of a voice input option of said voice input feature on said display, in case no such condition is fulfilled.

According to yet another aspect of the present invention a speech recognition system is provided that is capable of multimodal user interaction. The speech recognition system comprises at least  
30 one central processing unit, a display, a key-based input system, a microphone, and a data bus. Said display is connected to said central processing unit to be controlled by said central processing unit (CPU). Said key-based input system is operably connected to said central processing unit, to provide a key input feature providing key input options that can be displayed on said display. The microphone is operably connected to said at least one CPU to provide a  
35 audio-electronic converter to make voice input accessible to said CPU. The data bus is operably connected to said at least one CPU, to handle data and to exchange data required for the operation of the said at least one CPU.

BEST AVAILABLE COPY

Said at least one CPU comprises a first central processing unit and a second processing unit. Said first processing unit of said at least one CPU is configured to control multimodal interaction via said display, said key based input system and said microphone. Said first processing unit is further configured to monitor conditions that affect said voice input and to control and display an indication of a voice input option of said voice input feature on said display according to said condition. Said second central processing unit of said at least one CPU is configured to provide said voice input feature.

- 10 In another example embodiment of the present invention the first central processing unit and the second central processing unit of the at least one CPU are comprised in the same device.

In yet another example embodiment of the system the first central processing unit and the second central processing unit of the at least one CPU are comprised in different interconnected devices.

- 15 The interconnection can be provided by an audio telephone connection. The interconnection can be provided by a data connection such as GPRS (General Packet Radio Service), Internet, LAN (Local area network) and the like.

- 20 In another example embodiment said mobile electronic device further comprises a mobile telephone.

In the following, the invention will be described in detail by referring to the enclosed drawings in which:

- 25 Figure 1 is a flowchart of a method for dynamically indicating speech-enabling status to the user in multimodal mobile applications according to one aspect of the present invention,

Figure 2 is an example of an electronic device being capable of dynamically indicating speech-enabling status to the user for multimodal browsing, and

30

Figure 3 is an example of a display comprising different indications of visual input options and their actual possible input state, and

- 35 Figures 4A and 4B are examples of a distributed speech recognition system being capable of dynamically indicating speech-enabling status to the user for multimodal browsing.

Figure 1 is a flowchart of a method for dynamically indicating speech-enabling status to the user in multimodal mobile applications according to one aspect of the present invention. The method starts with the activation of a multimodal browsing 4. The expression 'multimodal browsing' is used to describe the possibility to interact with the device in different modes, i.e. the device can put out different modes e.g. a visual mode or an audible mode. Multimodal browsing can also include different input modes such as cursor or menu-keys and or alphanumerical keyboards, voice recognition or eye tracking. In the present figures a system with key and voice input capabilities is exemplary chosen to visualize the nature of the present invention. Following or simultaneously with the activation of the multimodal browsing, a monitoring or a surveying of the available input capabilities is started. The surveillance can be embodied by directly and repeatedly surveying the conditions that influence the speech recognition. The surveillance can also be embodied by a kind of indirect survey, by implementing sub algorithms at the respective application operating with a parameter that influences the speech recognition, and posts a signal or a message to the voice input application that a voice input is (probably) not possible. Such an approach can be described as an event based approach.

A possible condition is for example the actually available processing power. In case of a distributed voice input system a condition can be the connection properties such as bandwidth signal to noise ratio or the like. Another condition comprises the ambient or background noise which influences the speech recognition abilities.

From these example conditions it can be derived how probably a voice or speech input would be recognized. Therefore it can be derived, if the voice input feature is actually available or not. It should be noted that the ability to recognize certain voice inputs may vary from the condition. For example a background noise that comprises a sound signal that can be detected every second need not necessarily disturb the input of very short voice inputs, wherein voice inputs longer than a second can not be recognized because of the noise event.

In a next step a visual content is depicted 12 according to said monitored and evaluated input capabilities. That means that input options are depicted on a display of said electronic device or said mobile terminal device. Due to the usually restricted information content of a small mobile display, it should be clear that usually not all possible input options can be depicted on the display simultaneously. It should be noted that the unavailability of a voice input can also be depicted.

The user can simply perceive the available and possible speech inputs and can browse the elements depicted on the display by using either speech input or key input 16. When performing



multimodal browsing a new display content can be called and depicted, wherein the new content is also provided with speech input keywords or cues and the like which are dynamically generated by surveying and evaluating the multimodal browsing conditions (i.e. speech input / eye tracking / recognition conditions).

5

The method ends with the deactivation of the multimodal browsing 18. With the end of the multimodal browsing, the surveillance of the multimodal input conditions can also be stopped or interrupted. A direct connection between the boxes 8 or 12 to 18 has been economized, as the termination of the multimodal browsing is performed by a user input. In case of an automatic shutdown (e.g. a low battery power shutdown), the device can directly jump from 8 or 12 to 18.

10

As usability tests have indicated, the learning curve of users in using speech is steep in that users adopt the speech interaction rather quickly and fluently after the first successful attempts. However, there is a high threshold to overcome before the learning can start. In other words, users do not usually realize that speech input is available unless explicitly told so. Moreover, it takes time and courage for them to try the speech command if they are not sure about what they can say. After trial and success many start to even favor the speech input modality when making routine selections. After trial and error it may happen that users simply ignore any speech input ability.

20

The tasks where speech can be used in visual applications can be divided into two categories:

1) speech-enabling existing visual commands (selecting links, radio buttons, etc.)

2) allowing actions for which there is no visual equivalent (e.g. shortcuts = utterances combining several commands allowing the user to bypass hierarchical selections, or allowing the user to enter text, as in dictation)

25

This invention focuses mainly on category 1, indicating to the user what is speech-enabled and when at different points in the application. In category 2 type tasks, this invention allows indication to the user when speech input is possible, by selecting the implementation suitably, but what exactly can be said in these tasks is out of the scope of this invention, except in case of a combination with a speech input prediction system wherein the borders between the both categories become blurred.

30

To lower the threshold to use voice input and multimodal browsing a small demo version embodied in the electronic device or terminal can be embodied as some kind of a language lab, wherein the phone demonstrates in a replayed dialogue a typical input scenario with pre-recorded

35

BEST AVAILABLE COPY

speech inputs and input actions. For example : "To select the actual Battery status say 'say Fuelstate' repeat : ... .. and the requested information is read out loud 'Battery power at 25%'", "To select the actual Battery status say 'show Fuelstate' and the requested information is depicted on the display", wherein both actions can be accompanied the respective output.

5

In combination with a basic cursor based voice navigation system and speech recognizable words like 'right', 'left', 'up', 'down', 'click', 'doubleclick', 'clickclick', 'hold', 'delete' and 'select' a voice access can be provided even to voice-unable menu structures. The indication of a voice enabled speech navigation system can be provided by a mouth icon surrounded by the respective action icons or a mouth shaped cursor. In case of the selection of a gaming application by browsing via a menu (say „upupupupupclick“ or „game“) the possible speech input features are highlighted by a teeth / mouth icon or a snake icon to, select the game „snake“ (say „downdownclick“ or „snake“).

15 Figure 2 is an example of a electronic device or a terminal being capable of dynamically indicating speech-enabling status to the user for multimodal browsing. The device is depicted with a user interface as it is known from a mobile telephone. The mobile device is capable of executing multimodal interactive browsing, and comprises a user interface with input and output means such as the display 82 the keys 84 and 84', a microphone 86 and a loudspeaker 88. The user interface can be used for multimodal browsing comprising audio and key input and audio and display output. All the elements of the user interface are reconnected to a central processing unit CPU 80 to control the interaction of the user and the device.

25 The central processing unit is also connected to a data access means 90, to handle data and to exchange data required for the operation of the CPU 80 or applications running on said CPU 80. The CPU 80 is configured to perform multimodal browsing via said display (82), said key based input system 84, 84' and said microphone 86, and may be over said loudspeaker 88. The availability or operability of the multimodal browsing is dependent of parameters or on determined conditions. The CPU 80 can provide a multimodal browsing capability e.g. by running voice recognition applications on the device.

35 The CPU 80 is further connected to a data access means to access data stored in a built in storage (not shown) or access data via e.g. a network connection 92, to provide said multimodal browsing feature.

Said CPU 80 is further configured to monitor said conditions to continually determine the

availability of said voice input feature. The monitoring can be applied e.g. every second in shorter intervals or continuously, in dependence of the kind of parameters or the conditions that are monitored or surveyed.

- 5 The determined availability of the voice input feature is then visually indicated on a display on basis of said determined availability.

10 In the case that the multimodal browsing is constant independent from any external or internal restrictions the present invention can not be applied in a meaningful way, as if there are no changing parameters effecting the multimodal browsing, it is useless to monitor these parameters, as changes in vocabulary or the voice input capability can not occur.

15 Figure 3 is an example of a display comprising different indications of visual input options and their actual possible input state. There is depicted a display 58 of a mobile device that is multimodal browsing enabled. On the right side of the display 58 a light emitting diode LED 60 is placed. The LED can be used to indicate that a voice recognition engine or module is actually active or in a reception mode. The glowing flashing or blinking LED 60 can indicate that the user can talk to perform a user input or a user selection.

20 On the display there is depicted a usual list of selectable menu points "Menu option 1-4" 62. Related to each of the menu options 62 there is depicted an icon 64, 68 indicating the possible input modes. The "Menu options 1, 2 and 4" are provided with a mouth icon to indicate that these input options are "voice inputable". The "Menu option 3" is provided with a finger icon to indicate that the only available input option for this menu option is pressing a key.

25 The "Menu option 2" is underlined to indicate that a cursor is actually selectable by pressing an 'OK'-Button or by a voice input such as 'OK', 'click', 'doubleclick', 'clickclick' or 'select'.

30 The "Menu option 2" is depicted in bold letters to indicate that the "Menu option 2" is selectable by voice inputting the words " Menu option 2". The word 'option' of the "Menu option 1" is depicted in bold letters to indicate that the "Menu option 1" is selectable by voice inputting the words 'option'. The syllable 'men' and the number '4' of the "Menu option 4" are depicted in bold characters to indicate that the "Menu option 4" is selectable by voice inputting the words 'Men four', or a wording based on this abbreviation.

35

The icons 66, 70 on the bottom of the display 58 can also be used to indicate that a voice recognition engine or module is actually active or in or not in a reception mode. The icon 66, an open mouth can indicate that the user can talk to perform a user input or a user selection. The icons 70, closed lips sealed with a fingertip can indicate that the voice input option is actually not available.

The icons 66 and 70 and 64 and 68 can complement each other or exclude each other, as they provide redundant information.

- 10 Additionally to the icons, the following means can be used to denote when the user can speak:
- Spoken prompts can be played to the user, asking to speak an utterance ("Please choose / say a category.")
  - Playing an earcon (auditory icon, e.g. a beep) either alone or at the end of a prompt to indicate that the user can start speaking
  - 15 - The user can be allowed to control the speaking moment by clicking a special button to activate recognition (so called push-to-talk or "PTT" button)

In order to indicate what the user can say, the following means can additionally be used:

- Command lists are spoken to the user in the prompt ("Say 'Next', 'Previous', 'Back', 'Exit', or 'Help'")
- 20 - The prompt is designed to give implicit guidance to the user ("Do you want to go to Next or Previous?")
- The prompt gives an example about what can be said ("Select a day and a time, for example 'Monday at three'")

- 25 A spoken prompt is useful especially at the beginning of a session to remind the user about speech interaction. However, since the human beings can catch the content of a small mobile screen visually faster than it takes to listen to a sentence, prompts easily tend to sound long and tedious. Although barge-in (user interrupts the system prompt by speaking) is usually allowed in well-developed speech applications, users may be uncomfortable with speaking before the
- 30 system has stopped, since it is considered impolite in human-to-human conversations. A more serious problem with spoken prompts is that the information in them is usually lost beyond recovery if the user is not concentrating. Also, long command lists are not useful, since they increase the user's memory load and boredom, since nearly every computer generated monologue

lasting longer than 7 words or 3 seconds can readily be perceived as boring or annoying.

To summarize, while prompts are useful in making the situation more dialogue-like, they tend to be too long and available only for a short time. Auditory icons are short but they are also temporary signals. Visual cues for speaking that would stay visible on the screen to indicate when speech is allowed, when it is not, and what exactly can be said, would be an easy and transparent way to indicate speech-enabling to the user. Indicating when speech is allowed is also an easy way to make users aware of the barge-in feature and encourage them to interrupt or "vocally override" possible prompts.

Push-to-talk buttons, while allowing the user more control of the interaction, are not fully without problems, either. The device has to have a separate button for voice activation, or the user must be separately taught that a button serves as a push-to-talk button in some contexts. In some mobile contexts, pressing even one button might be cumbersome e.g. while riding on a motorbike on the pillion.

Figure 4A and 4B are examples of a distributed speech recognition system being capable of dynamically indicating speech-enabling status to the user for multimodal browsing.

Figure 4A is an example of a distributed speech recognition system being capable of dynamically indicating speech-enabling status to the user for multimodal browsing, wherein said distributed speech recognition system is integrated in a single device 77. The term "distributed speech recognition" is used to indicate that the multimodal browsing and the speech recognition is executed at least in different processing units of said single device 77.

The mobile device 77 is comprises of a speech recognition system that is capable of executing multimodal interactive browsing, and comprises a user interface with input and output means such as the display 82 the keys 84 and 84', a microphone 86 and a loudspeaker 88. The user interface can be used for multimodal browsing comprising audio and key input and audio and display output. All the elements of the user interface are reconnected to a central processing unit CPU 80 to control the interaction of the user and the device.

The speech recognition system comprises at least one central processing unit 80, a display 82, a key-based input system 84, 84', a microphone 86, and a data bus 91. Said display is connected to said central processing unit to be controlled by said CPU 80. Said key-based input system 84, 84' is operably connected to said central processing unit 80, to provide a key input feature providing key input options that can be displayed on said display 82.

The microphone 86 is operably connected to said at least one CPU 80 to provide a audio-electronic converter to make voice input accessible to said CPU 80. The data bus 91 is operably connected to said at least one CPU 80, to handle data and to exchange data required for the operation of the said at least one CPU 80. The data bus 91 is operably connecting said at least one CPU 80 to an internal memory 83 to provide a data access to stored data necessary to provide said key input feature and/or said voice input feature. The internal memory 83 can e.g. store the different condition and combinations of conditions of the device in which the voice input feature is accessible or not.

Said at least one CPU 80 comprises a first central processing unit 81 and a second processing unit 81'. Said first processing unit 81 of said at least one CPU 80 is configured to control multimodal interaction via said display 82, said key based input system 84, 84' and said microphone 86. Said first processing unit 81 is further configured to monitor conditions that affect said voice input and to control and display an indication of a voice input option of said voice input feature on said display 82 according to said monitored condition.

Figure 4B is an example of a distributed speech recognition system being capable of dynamically indicating speech-enabling status to the user for multimodal browsing that is distributed between at least two devices. A distributed voice recognition can comprise the advantages that the resources required for speech recognition can be economized in the small and e.g. portable device 78.

To provide a distributed system, the CPU 80 has to be distributed between the two devices. The first central processing unit 81 and the second central processing unit 81' of the at least one CPU 80 are comprised in different interconnected devices 78 and 79. The interconnection between the two devices (and of course the first central processing unit 81 and the second central processing unit 81') can be provided by, e.g., a telephone connection. The interconnection can also be provided by a data connection such as GPRS (General Packet Radio Service), Internet, LAN (Local Area Network) and the like.

Said first central processing unit 81 alone can be configured to monitor said conditions to continually determine the availability of said voice input feature. The monitoring can be applied e.g. every second in shorter intervals or continuously, in dependence of the kind of parameters or the conditions that are monitored or surveyed.

The major advantage of the invention is that it can be applied to any kind of mobile electronic

devices regardless of the used features. A user using an electronic device always under the best voice control or multimodal browsing conditions will not recognize the presence of the present invention. The present invention can be applied to any kind of voice control or voice input used in technical applications. There is also a possibility to apply the present invention to a non mobile system with no limitations in regard of resources. In a non mobile system the present invention can be used to indicate the words that can be recognized with a probability of nearly 100% and words that can be recognized only with a lower recognition rate and therefore are not to be regarded as being available (or requiring more training).

- 10 The visual keyword or cue that is chosen to mark the speech-enabling could be a color scheme or some other method, such as underlining. Underlining might easily be confused with a hyperlink, however. Color would be a good choice, and color displays are becoming more and more general. Red is typically used to mark active recording in audio applications, so it might be a suitable choice to indicate that speech-enabling is on. Some traffic light scenario could also be adopted. Animated icons may help to visualize that a longer action e.g. a voice input is possible for a depicted element such as ant colons, an animated sound spectrum monitor a talking mouth.

- 20 The color system must be learnt as well, even if only two colors are used, one for speech-on and the other for speech-off indications. A small legend describing the color usage might be visible on the early screens of the application.

Instead of colors, the speech-enabled commands could be marked in some other way, e.g. drawing a small speech bubble around the command. The visual cue should be directly tied to the command, however, to make the enabling method as transparent to the user as possible.

25

Changing the visual cue dynamically while on the same page can be done with suitable event mechanism. In the same way as the browser can highlight visual symbols in an XHTML application when a suitable 'onclick' or 'onfocus' event is caught, new events can be defined for cases that call for change in the visual speech-enabling cue. When a multimodal mobile browser catches these events, it would then change the color or other chosen visual cue in corresponding GUI elements as required.

30

With speech-enabled tasks that have no visual equivalent, some traffic light scheme can be used to indicate when speech recognition is active or inactive. This is relatively easy to implement

with events that affect the whole screen at a time. One such measure can be to wobble the display illumination, invert the depiction mode, or selectively animating the voice enabled menu points, or let small balls jump from syllable to syllable as known from 'Karaoke' videos.

- 5 Additional features that can be combined with the present invention are e.g. input prediction, training dialogues, voice input proposals via text or speech output. Icon based menu structures for illiterate people, trainable speech input. Read out user manuals employing a „read out“ and a „read in“ key.
- 10 This application contains the description of implementations and embodiments of the present invention with the help of examples. It will be appreciated by a person skilled in the art that the present invention is not restricted to details of the embodiments presented above, and that the invention can also be implemented in another form without deviating from the characteristics of the invention. The embodiments presented above should be considered illustrative, but not
- 15 restricting. Thus the possibilities of implementing and using the invention are only restricted by the enclosed claims. Consequently various options of implementing the invention as determined by the claims, including equivalent implementations, also belong to the scope of the invention.



### Claims

- 5 1. Method for indicating speech-enabled input for multimodal interaction in an electronic device having a user interface, comprising:  
activating a multimodal user interaction feature of said user interface in which at least one key input option and at least one voice input option is provided,  
displaying the at least one key input option on a display of said electronic device,  
characterized by  
10 checking, if at least one condition generally affecting voice input is fulfilled, and  
providing said at least one voice input option and displaying indications of said voice input options on said display according to said condition.
- 15 2. Method according to claim 1, wherein said displayed indications of voice input options comprise keywords.
- 20 3. Method according to claim 1 or 2, wherein said displaying of indications of said voice input options on said display further comprises displaying if a speech recognition is actually possible.
4. Method according to any of the preceding claims, wherein said displaying of indications of voice input options comprises displaying said voice input options.
- 25 5. Method according to any of the preceding claims, wherein said displaying of indications of said voice input options on said display, is provided with a hysteresis.
6. Method according to any of the preceding claims, wherein said displaying of indications of said voice input options on said display is provided with a backlog function.
- 30 7. Software tool comprising program code means stored on a computer readable medium for carrying out the method of anyone of claims 1 to 6, when said software tool is run on a computer or network device.
- 35 8. Computer program product comprising program code means stored on a computer readable medium for carrying out the method of anyone of claims 1 to 6, when said program product is run on a computer or network device.

9. Computer program product comprising program code, downloadable from a server for carrying out the method of anyone of claims 1 to 6, when said program product is run on a computer or network device.

5 10. An electronic device capable of executing multimodal interactive browsing, comprising:  
a central processing unit CPU (80),  
a display (82) connected to said CPU (80), to display visual content received from said CPU (80) on said display (82),  
a key-based input system (84, 84') operably connected to said CPU (80), to provide a key  
10 input feature providing key input options displayed on said display ,  
a microphone (86) operably connected to said CPU (80), to provide a voice input feature, and  
a data bus (90), operably connected to said CPU (80), to handle data and to exchange data required for the operation of the CPU (80),  
wherein said CPU (80) is configured to control multimodal interaction via said display (82),  
15 said key based input system (84, 84') and said microphone (86), and  
wherein said CPU (80) is configured to monitor conditions that affect said voice input, and to provide said voice input feature and display an indication of a voice input option of said voice input feature on said display (82) according to said condition.

20 11. Electronic device according to claim 10, further comprising a mobile communication device.

12. A speech recognition system capable of multimodal interaction and having a user interface, comprising:  
25 at least one central processing unit CPU (80),  
a display (82) connected to said CPU (80),  
a key-based input system (84, 84') operably connected to said CPU (80), to provide a key input feature providing key input options displayed on said display,  
a microphone (86) operably connected to said at least one central processing unit (80),  
30 a data bus (91), operably connected to said at least one CPU (80), to handle data and to exchange data required for the operation of the said at least one CPU (80)  
wherein a first central processing unit (81) of said at least one CPU (80) is configured to control multimodal interaction via said display (82), said key based input system (84, 84') and said microphone (86) and to monitor conditions that affect said voice input and to control  
35 and display an indication of a voice input option of said voice input feature on said display (82) according to said condition, and  
wherein a second central processing unit (81') of said at least one CPU (80) is configured to

provide said voice input feature.

13. A system according to claim 12, wherein the first central processing unit (81) and the second central processing unit (81') are comprised in the same device (77).

5

14. A system according to claim 12, wherein the first central processing unit (81) and the second central processing unit (81') are comprised in different interconnected devices (78, 79).

**Abstract of the disclosure**

The present invention provides a method, a device and a system for multimodal interactions. The  
5 method according to the invention comprises the steps of activating a multimodal user  
interaction, providing at least one key input option and at least one voice input option, displaying  
the at least one key input option, checking if there is at least one condition affecting said voice  
input option, and providing voice input options and displaying indications of said provided voice  
10 input options according to said condition. The method is characterized by checking if at least one  
condition affecting the voice input is fulfilled and providing said at least one voice input option  
and displaying indications of said voice input options on said display, according to said  
condition.

1/2

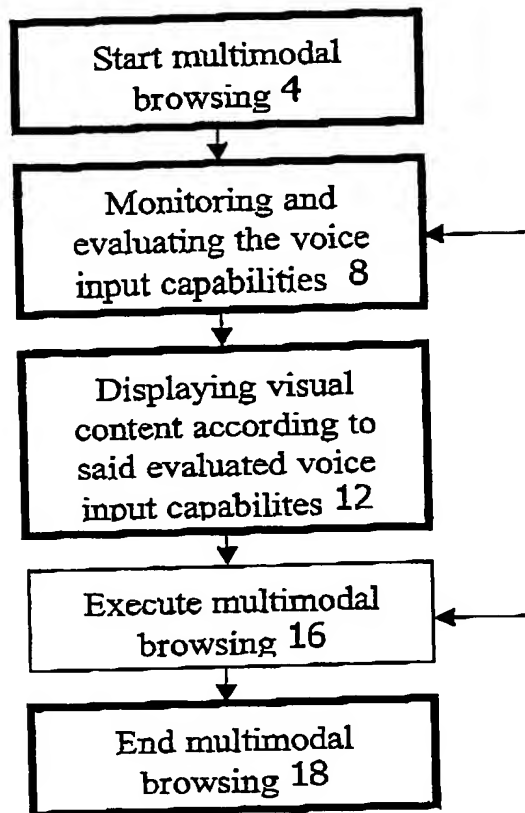


Fig. 1

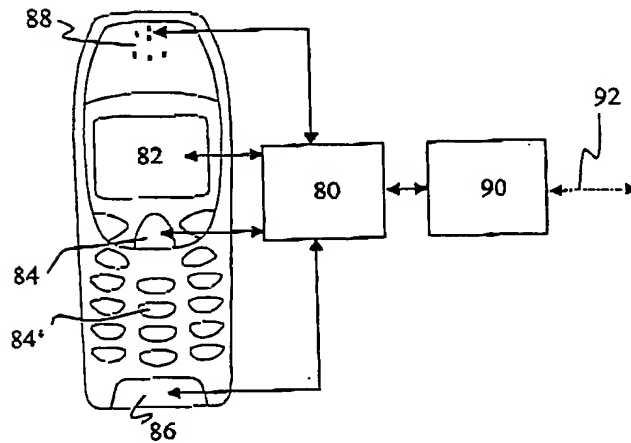


Fig. 2

2/2

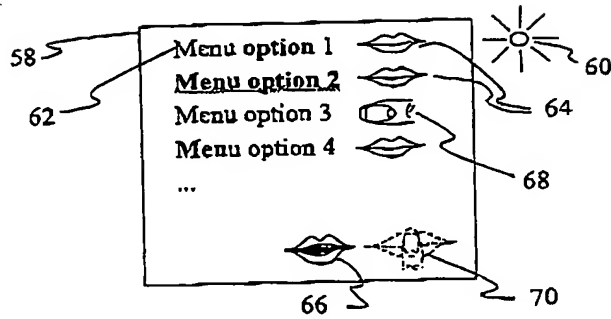


Fig. 3

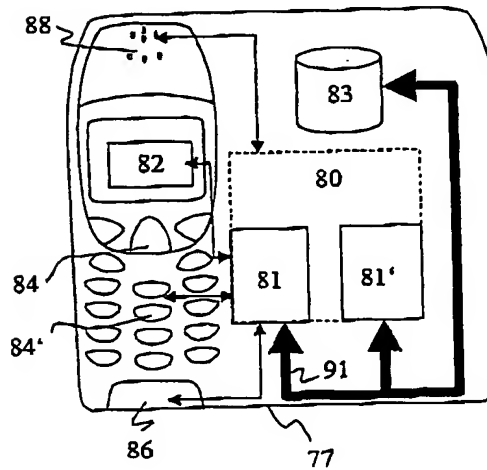


Fig. 4A

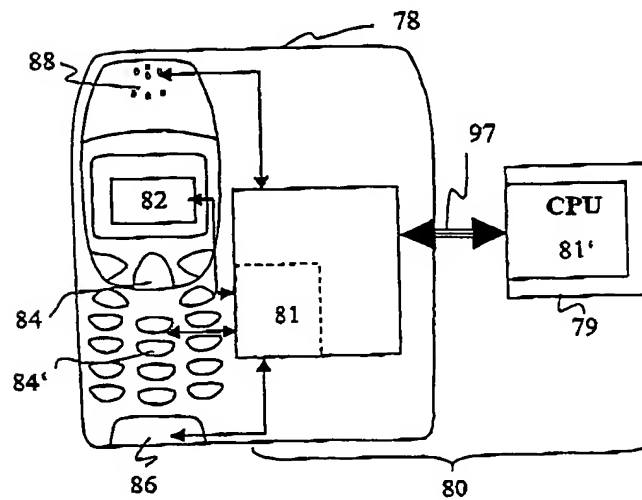


Fig. 4B